ABSTRACT OF THE DISCLOSURE

Disclosed herein are systems and methods for information organization whereby characteristics regarding entities are inferred from the characteristics of similar entities. This is referred to herein as a "fuzzy similarity join" and is exemplified using a chemical similarity join. The disclosed systems and methods are particularly useful for chemical compound analysis in the area of pharmaceutical development.